

# PENYUSUNAN INSTRUMEN TES FORMATIF FISIKA SMP

<sup>1)</sup>Isnaini Rohayati, <sup>2)</sup>Dr.Nonoh Siti A, M.Pd, <sup>3)</sup>Elvin Yusliana E, S.Pd, M.Pd

Fakultas Keguruan dan Ilmu pendidikan  
Universitas Sebelas Maret

Jl. Ir. Sutami 36 A, Surakarta, email : [ir.ice\\_na@yahoo.co.id](mailto:ir.ice_na@yahoo.co.id)

## ABSTRACT

The aim of this research was to apprehend the characteristics of physics formative test instrument which has been arranged.

This research was a basic research. The research had been done in Surakarta and Girimarto involved respondent of students of class VIII A, B, C, D, E, F, and G SMP Negeri Girimarto on educational year 2011/2012. The test that had been arranged was a formative test of light matter. Data gathering was done by non-test technique and test technique. Data analysis technique that was used was qualitative analysis with problem study and quantitative technique with multiple choices problem analysis covered difference capacity, difficulty level, distractor effectivity, and problem reliability.

The research finding was a qualitative study result by expert. From the qualitative study, it was obtained that many problems need to be revised since the answer choice were not alphabetical, ineffective sentences, and inappropriate matter. The result of field test I showed as follow: packet A: there were 6 problems of difficult category, 24 problems of medium category, and 10 complicated problems, from the difference capacity, there were 5 problems that had negative difference capacity, 5 problems were poor, 5 problems were satisfactory, 13 problems were good, and 12 problems were excellent. For the packet B problems, there were 13 difficult problems, 23 problems were medium, and 4 problems were easy. There were 12 problems rejected, 1 problems was poor, 4 problems were satisfactory, 15 problems were good and 10 problems were excellent based on the difference capacity.

From the research finding, it could be concluded that the arranged test was “tes Formatif Fisika SMP Materi Cahaya, Tahun Pelajaran 2011/2012”. The test was an objective test in multiple choices form with four answer choices of 50 problems. Time allocation of test doing was 90 minutes. The arranged test was equipped by direction how to do and key answer. Characteristics of arranged test were packet A problems with 25% easy category, 60% medium category, and 15% difficult category. From the difference capacity, it was obtained that problems with excellent difference capacity were 30%, good were 32,5%, satisfactory were 12,5%, poor were 12,5%, and rejected were 12,5%. From the effectiveness distractor aspect, it was obtained that 85% of problems were four distractors functioned, and 15% problems were 3 distractors functioned, 0% problems were two distractors functioned, and 0% problems were one distractors functioned. Packet B problems, there were 10% of easy category, 57,5% of medium category, and 32,5% of difficult category. Based on the difference capacity, the problems with excellent were 25%, good were 32,5%, satisfactory were 10%, poor were 2,5%, and rejected were 30%. Whereas from the effectiveness distractors aspect, it was acquired that 85% of problems were four distractors functioned, and 12,5% problems were 3 distractors functioned, 2,5% problems were two distractors functioned, and 0% problems were one distractor functioned. From the last test instrument, it was attained that there were 22% of easy category, 54% of medium category, and 24% difficult category. Problems with excellent difference capacity were 10%, good were 36%, satisfactory were 28%, poor were 18%, and rejected were 8%. From the effectiveness distractor aspect, it was gained that 68% problems were four distractors functioned, and 28% problems were 3 distractors functioned, 4% problems were two distractors functioned, and 0% problems were one distractor functioned. From the three packet of test, the test that had the best quality and characteristics was the last test.

**Key words:** Test arrangement, Formative test, Cognitive aspect.

## PENDAHULUAN

Kegiatan belajar mengajar di kelas tidak terlepas dari interaksi antara pendidik dan peserta didik, dilengkapi dengan sarana prasarana penunjang dan diakhiri dengan evaluasi (penilaian). Penilaian merupakan unsur yang penting dalam kegiatan belajar mengajar. Penilaian merupakan rangkaian kegiatan untuk memperoleh, menganalisis dan menafsirkan data tentang proses dan hasil belajar peserta didik yang dilakukan secara sistematis dan berkesinambungan sehingga menjadi informasi yang bermakna dalam pengambilan keputusan (Sri Yamtinah : 2009:1). Penilaian bukan semata – mata untuk mengelompokkan siswa pada kelas tertentu, memberikan predikat bagi yang “unggul” dan bukan pula sebagai ajang labelisasi kemampuan siswa. Mengingat akibat dari pemberian label itu sendiri dapat merusak motivasi belajar dan sangat merugikan, karena penilaian sendiri memiliki tujuan yang lebih penting, yakni mengetahui tingkat kompetensi peserta didik, mengukur pertumbuhan dan perkembangan peserta didik, mendiagnosis kesulitan belajar siswa, dan mengetahui pencapaian kurikulum.

Terlepas dari esensi dan urgensi tes sebagai suatu metode untuk menilai, fakta yang didapatkan di sekolah adalah guru tidak mempersiapkan rancangan tesnya dengan maksimal, kebanyakan soal tes diambil dari LKS atau buku pegangan guru. Selain itu, soal tes yang digunakan tidak diketahui standardisasinya.

Kenyataan ini tidak sesuai dengan Permendiknas No 16 tahun 2007 mengenai kualifikasi akademik dan Standar Kompetensi Guru khususnya kompetensi pertama yang berbunyi “memahami prinsip – prinsip penilaian dan evaluasi proses serta hasil belajar sesuai dengan karakteristik mata pelajaran yang diampu” dan kompetensi keempat “mengembangkan instrumen penilaian dan evaluasi proses dan hasil belajar. Jika instrumen yang digunakan guru untuk menilai hanya diambil dari LKS atau buku pegangan dan soal tidak diketahui standar-disasinya, bisa jadi soal yang digunakan terlalu susah atau terlalu mudah, sehingga hasil tes tidak memberikan gambaran yang representatif tentang kompetensi siswa yang telah atau belum dimiliki oleh siswa.

Apabila tes sudah dipersiapkan dengan optimal dan instrumen yang berkualitas baik, maka informasi yang diperoleh dari hasil tes dapat menunjukkan sejauh mana kompetensi yang ditetapkan telah tercapai. Sehingga informasi tersebut dapat dijadikan balikan untuk meningkatkan kualitas dan menyempurnakan proses pembelajaran.

Penelitian ini berjudul **PENYUSUNAN INSTRUMEN TES FORMATIF FISIKA UNTUK SMP**. Dengan disusunnya instrumen tes formatif yang baik diharapkan dapat membantu guru untuk mengetahui tingkat kemampuan siswa dengan tepat sehingga informasi hasil penilaian dapat digunakan guru untuk mengetahui tingkat keberhasilan pelaksanaan proses pembelajaran dan acuan untuk mengambil keputusan.

Dari latar belakang masalah yang telah diuraikan di atas maka dapat diidentifikasi permasalahan sebagai berikut (1). Instrumen tes yang digunakan oleh guru diambil dari LKS dan buku pegangan, (2) Tes yang digunakan oleh guru belum diteliti standardisasinya.

Agar dalam pembahasan permasalahan ini lebih mendalam dan cakupannya tidak terlalu luas maka permasalahan-permasalahan yang ada dibatasi sebagai berikut (1). Tes yang disusun merupakan tes formatif Fisika SMP kelas VIII materi Cahaya, (2) Jenis tagihan yang akan dibuat berupa tes tertulis berbentuk soal pilihan ganda khusus untuk mengukur aspek kognitif.

Berdasarkan latar belakang masalah, identifikasi masalah dan pembatasan masalah yang telah dikemukakan, maka dapat dirumuskan permasalahan sebagai berikut ; Bagaimanakah karakteristik instrumen tes formatif Fisika untuk SMP kelas VIII yang telah disusun?

Tujuan yang ingin dicapai dalam penelitian ini adalah mengetahui karakteristik instrumen tes formatif fisika yang telah disusun

Dari hasil penelitian ini diharapkan dapat memberikan manfaat antara lain (1) Memberikan informasi tentang instrumen tes formatif fisika untuk tingkatan SMP, (2) Memberikan informasi tentang cara penyusunan instrumen tes yang baik, (3) Sebagai bahan pertimbangan dan masukan atau re-ferensi ilmiah untuk penelitian lebih lanjut.

Asumsi dan keterbatasan dalam penelitian ini antara lain (1). Instrumen tes yang telah disusun diasumsikan layak untuk diteskan karena tes disusun berdasarkan prosedur dan memenuhi kriteria penyusunan tes, (2). Bentuk soal pilihan ganda memungkinkan peserta tes untuk menebak jawaban.

## KAJIAN TEORI

Ada dua istilah yang terkait dengan konsep penilaian (*assessment*), yaitu pengukuran (*measurement*) dan evaluasi (*evaluation*). (Djemari Mardaphi, 2007:1). Menurut Griffin dan Nix “pengukuran, asesmen dan evaluasi adalah hirarki”. Evaluasi di-dahului dengan penilaian (*assessment*) sedangkan penilaian didahului dengan pengukuran. Pengukuran membandingkan hasil pengamatan dengan kriteria, asesmen menjelaskan dan menafsirkan hasil pengukuran, sedangkan evaluasi adalah penetapan nilai atau implikasi suatu perilaku (Djemari Mardapi, 2008:1). Penilaian hasil belajar peserta didik harus memperhatikan kesahihan, objektivitas, adil, keterpaduan, keterbukaan, kesinambungan, aturan kriteria dan akuntabilitas (Yamtinah, 2009:1).

Sebuah penilaian pembelajaran dilakukan dengan beberapa tujuan. Menurut Mundilarto, penilaian bertujuan untuk “mengetahui tingkat kompetensi peserta didik, mengukur pertumbuhan dan perkembangan, mendiagnosis kesulitan belajar, mengetahui hasil proses belajar mengajar, mengetahui pencapaian kurikulum, mendorong guru untuk mengajar lebih baik, sebagai upaya meningkatkan akuntabilitas lembaga dan sebagai upaya meningkatkan kualitas pendidikan” (2010 : 16).

Untuk mendapatkan informasi mengenai ukuran kemampuan peserta didik maka diperlukan suatu alat ukur. Alat ukur yang dipakai dalam bidang evaluasi biasanya berupa tes dan non tes. Istilah tes, pengukuran dan evaluasi merupakan istilah umum dijumpai dalam kegiatan evaluasi. Djemari berpendapat bahwa “tes merupakan salah satu cara untuk menaksir besarnya tingkat kemampuan atau keterampilan seseorang secara tidak langsung, yakni melalui respon orang tersebut terhadap sejumlah stimulus atau pertanyaan” (2008 : 67).

Bentuk tes yang biasa digunakan ada beberapa macam tergantung penggunaan. “Bentuk tes yang digunakan di lembaga pendidikan dapat dikategorikan menjadi dua jenis, yaitu tes objektif dan tes non objektif” (Djemari Mardapi, 2008: 69). Bentuk tes objektif yang sering digunakan adalah bentuk pilihan ganda, menjodohkan, benar salah, dan uraian objektif

Formatif berasal dari kata *form*. Secara umum tes ini berfungsi untuk memperoleh informasi atau masukan tentang tingkat keberhasilan pelaksanaan proses pembelajaran. Masukan ini berguna untuk memperbaiki strategi mengajar. Secara ringkas, tes formatif memiliki karakteristik sebagai berikut: 1) dilakukan pada akhir setiap satuan pelajaran, 2) bertujuan untuk mengetahui sejauh mana tujuan pembelajaran pada setiap satuan pelajaran sudah

tercapai, 3) berfungsi memberikan umpan balik untuk perbaikan proses belajar mengajar, 4) dilakukan dengan mempergunakan tes hasil belajar, kuesioner ataupun cara lain yang sesuai, 5) peserta didik dinilai berhasil jika mencapai taraf penguasaan minimal 70 % dari tujuan pembelajaran yang ingin dicapai. (Mundilarto, 2010 : 49).

Penyusunan instrumen tes melalui beberapa tahapan, yakni "(1) melakukan analisis kurikulum, (2) menetapkan tujuan tes, (3) membuat kisi (4) menulis instrumen tes, (5) melakukan telaah instrumen secara teoritis, (6) melakukan uji coba dan analisis item hasil uji coba tes, dan (7) merevisi instrumen tes" (Estina Ekawati, 2011:20).

Setelah tes disusun, tes selanjutnya dianalisis. Analisis hasil tes dapat dilakukan secara kualitatif dan kuantitatif. Analisis hasil tes dilakukan untuk melihat secara tidak langsung kualitas dari tes yang telah disusun. Analisis kualitatif dilakukan dengan penelaahan pada tes yang telah dibuat. Telaah tes secara kualitatif dilakukan terhadap aspek materi, konstruksi dan bahasa. Djemari (2008:137) memberikan definisi mengenai aspek materi, konstruksi dan bahasa sebagai berikut "Aspek materi berkaitan dengan substansi keilmuan yang ditanyakan serta tingkat berfikir yang terlibat. Aspek konstruksi berkaitan dengan teknik penulisan soal, baik bentuk objektif maupun yang non objektif. Aspek bahasa terkait dengan kekomunikatifan / kejelasan hal yang ditanyakan.

Analisis kuantitatif yang dilakukan meliputi analisis mengenai validitas, reliabilitas, taraf kesukaran, dan efektifitas pengecoh. Sebuah tes dapat dikatakan valid apabila tes dapat tepat mengukur apa yang hendak diukur. Valid dapat diartikan sebagai sah, sehingga validitas dapat diartikan sebagai kesahihan. Walaupun istilah "tepat" belum tentu mencakup semua arti yang tersirat dalam kata "valid" dan kata "tepat" kadang – kadang digunakan dalam konteks yang lain, akan tetapi tambahan kata "tepat" dalam menerangkan kata "valid" dapat memperjelas apa yang dimaksud.

Sebuah tes dapat dikatakan reliabel jika memberikan hasil yang tetap apabila diteskan berkali – kali. Dengan kata lain jika kepada siswa diberikan tes yang sama pada waktu yang berlainan, maka setiap siswa akan tetap berada dalam urutan yang sama dalam kelompoknya.

Syaifuddin mengemukakan bahwa "pendekatan konsistensi internal dimaksudkan untuk menghindari masalah yang biasanya ditimbulkan oleh pendekatan tes ulang dan bentuk paralel. Pendekatan konsistensi internal, hanya diperlukan satu kali tes kepada sekelompok individu. Oleh karena itu pendekatan ini memiliki nilai praktis dan efisiensi yang tinggi" (1997:63).

Berkaitan dengan penelitian ini, koefisien reliabilitas dicari dengan menggunakan teknik konsistensi internal metode koefisien alpha. Nilai koefisien reliabilitas dicari dengan menggunakan bantuan *software Microcat ITEMAN 3,0*.

Persamaan matematis untuk mencari koefisien alpha adalah

$$\alpha = \left[ \frac{k}{k-1} \right] \left[ 1 - \frac{\sum s_j^2}{s_x^2} \right]$$

(Syafuddin Azwar, 1997 :77)

keterangan

k : banyaknya belahan tes

$s_j^2$  : varians belahan j

$s_x$  : varians skor tes

kriteria

$0,00 \leq r \leq 0,19$  : reliabilitas sangat rendah

$0,20 \leq r \leq 0,39$  : reliabilitas rendah

$0,40 \leq r \leq 0,59$  : reliabilitas cukup

$0,60 \leq r \leq 0,79$  : reliabilitas tinggi

$0,80 \leq r \leq 1,00$  : reliabilitas sangat tinggi

(Suharsimi Arikunto, 2005 : 100)

Analisis kuantitatif yang lain adalah analisis mengenai taraf kesukaran. Taraf kesukaran tes adalah kemampuan tes tersebut dalam menjangkau banyaknya subyek peserta tes yang dapat mengerjakan dengan betul. Rumus yang digunakan untuk menghitung indeks kesukaran adalah :

$$P = \frac{B}{JS}$$

( Suharsimi Arikunto, 2001 : 208 )

dengan

P : indeks kesukaran

B :banyaknya siswa yang menjawab benar

JS : jumlah siswa peserta tes

klasifikasi soal :

$1,00 \leq p \leq 0,29$  ;soal kategori sukar

$0,30 \leq p \leq 0,69$  ; soal kategori sedang

$0,70 \leq p \leq 1,00$  ; soal kategori mudah

Untuk mengetahui keefektifan soal dalam membedakan peserta tes kelompok atas dan kelompok bawah, maka dilakukan analisis daya pembeda. Daya pembeda butir soal berfungsi untuk menentukan dapat tidaknya suatu butir soal membedakan kelompok peserta tes yang berkemampuan tinggi dan berkemampuan rendah. Daya pembeda dapat dihitung dengan mencari nilai korelasi Point Biserial dan nilai korelasi Biserial. Secara matematis nilainya dinyatakan dengan:

$$r_{pbis} = \frac{M_p - M_t}{S_T} \sqrt{\frac{p}{1-p}} = \frac{M_p - M_t}{S_T} \sqrt{\frac{p}{q}}$$

(Nonoh Siti Aminah, 2012 :9)

keterangan

p : jumlah yang menjawab benar

M<sub>p</sub> : mean skor tes dari peserta tes yang menjawab benar

M<sub>t</sub> : mean skor total

S<sub>T</sub> : varians total

Dan nilai korelasi biserial dinyatakan dengan persamaan

$$r_{bis} = \left( \frac{M_p - M_t}{S_T} \right) \left( \frac{p}{y} \right)$$

Nilai korelasi Biserial selalu lebih rendah dibanding dengan korelasi Point Biserial. Klasifikasi daya beda butir soal adalah sebagai berikut.

Tabel 2.1. Klasifikasi Daya Beda

Nilai p	Klasifikasi
$p \leq 0,19$	<i>Poor</i>
$0,20 \leq p \leq 0,40$	<i>Satisfactory</i>
$0,40 \leq p \leq 0,70$	<i>Good</i>
$0,70 \leq p \leq 1,00$	<i>Excellent</i>
Bertanda negatif	ditolak

Sebagai tindak lanjut dari penganalisisan mengenai daya pembeda item, item soal dikelompokkan dengan dasar berikut :

1. Butir – butir soal yang sudah memiliki daya pembeda item yang baik (*satisfactory*, *good*, dan *excellent*) hendaknya dimasukkan dalam buku bank soal tes hasil belajar. Butir – butir item tersebut dapat dipakai.
2. Butir – butir item yang daya pembedanya masih rendah (*poor*), ada dua kemungkinan tindak lanjut, yaitu :
  - a) Ditelusuri untuk kemudian di-perbaiki, dan setelah diperbaiki dapat diajukan lagi dalam tes hasil belajar yang akan datang.
  - b) Dibuang dan tidak digunakan.

(Anas Sudijono, 1995 :389)

Analisis efektifitas pengecoh atau analisis distribusi jawaban dimaksudkan untuk mengetahui apakah alternatif jawaban yang tersedia sudah berfungsi dengan baik atau belum. Pengecoh dikatakan berfungsi dengan baik apabila jumlah pemilihnya paling sedikit 2% dari jumlah peserta tes.

## METODOLOGI PENELITIAN

Penelitian dilaksanakan di SMP Negeri 2 Girimarto kelas VIII tahun pelajaran 2011/2012. Penelitian dilaksanakan pada bulan Januari – Juni 2012. Subjek penelitian adalah siswa kelas VIII A, B, C, E, F dan G SMP Negeri 2 Girimarto tahun pelajaran 2011 / 2012.

Metode penelitian yang digunakan merupakan metode penelitian dasar. “Penelitian dasar adalah suatu pene-litian yang bertujuan untuk mengembangkan teori ilmiah atau prinsip dasar suatu disiplin ilmu yang lebih baik daripada hanya me-mecahkan persoalan praktis” ( Nonoh Siti A, 2002:1).

Analisis data dilakukan secara kuantitatif dan kualitatif. Analisis data secara kualitatif dilakukan untuk menganalisis butir soal yang telah disusun secara deskriptif, Penelaahan dilakukan meliputi aspek materi, konstruksi dan bahasa. Dalam penelaahan soal digunakan lembar telaah untuk mempermudah prosedur pelaksanaan. Adapun analisis kuantitatif yang dilakukan meliputi analisis validitas, tingkat kesukaran item soal, daya pembeda dan efektifitas disktraktor.

## HASIL

Sebelum instrument diujikan ke lapangan, instrumen ditelaah secara kualitatif. Telaah instrumen secara teoritis atau deskriptif dilakukan untuk melihat keterbacaan instrumen dan untuk validasi isi. Telaah instrumen secara deskriptif dilakukan dengan peninjauan aspek materi, konstruksi dan bahasa. Dalam penelitian ini, telaah instrumen secara deskriptif dilakukan oleh pembimbing, dosen ahli materi dan guru kelas. Hasil telaah deskriptif ditemukan beberapa soal yang belum sesuai dengan kriteria sehingga harus direvisi, misalnya

pilihan jawaban tidak alfabetis, pilihan jawaban mengulang kata, soal tidak efektif dan terdapat materi yang tidak diajarkan di sekolah.

Setelah instrumen tes ditelaah secara deskriptif, selanjutnya instrumen tersebut diujicobakan pada siswa kelas VIII B, VIII C dan VIII E SMP Negeri 2 Girimarto. Uji coba dilaksanakan pada Rabu, 16 Mei 2012. Hasil uji coba selanjutnya dianalisis menggunakan *Microcat ITEMAN 3.0* untuk mengetahui nilai taraf kesukaran, daya pembeda, indeks reliabilitas dan efektifitas distraktor. Berdasarkan analisis hasil uji I pada 47 responden untuk setiap paket soal, untuk paket soal A diperoleh 6 soal (9, 24, 25, 29, 30, 38) sukar, 24 soal (4, 5, 6, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 26, 27, 28, 32, 33, 34, 36, 37, 39) kategori sedang dan 10 soal (1, 2, 3, 7, 8, 10, 14, 31, 35, 40) yang mudah. Sedangkan untuk paket soal B diperoleh 13 soal (1, 3, 6, 9, 20, 23, 24, 25, 26, 29, 30, 37, 38) sukar 23 soal (2, 4, 5, 10, 11, 13, 14, 15, 16, 17, 18, 19, 21, 22, 27, 28, 31, 32, 33, 34, 36, 39, 40) termasuk kategori sedang dan 4 soal (7, 8, 12, 35) yang mudah.

Analisis kuantitatif yang lain adalah analisis mengenai daya pembeda soal. Analisis daya beda soal bertujuan untuk mengetahui dapat tidaknya suatu soal membedakan kelompok atas dan kelompok bawah. Indeks daya pembeda didapatkan dari selisih proporsi yang menjawab dari masing – masing kelompok. Dari hasil analisis ITEMAN untuk kedua paket soal, pada paket soal A diperoleh 5 soal (1, 9, 24, 31, 38) memiliki daya pembeda *poor*, 5 soal (3, 4, 5, 7, 8) *satisfactory*, 13 soal (2, 6, 10, 11, 12, 14, 17, 20, 26, 35, 36, 37, 40) *good*, 12 soal (soal nomor 13, 15, 16, 18, 19, 21, 22, 23, 30, 33, 34, 39) *excellent* dan 5 soal (soal nomor 25, 27, 28, 29, 32) ditolak. Sedangkan untuk paket soal B, diperoleh 1 soal (soal nomor 7) mempunyai daya pembeda *poor*, 4 soal (soal nomor 1, 2, 20, 21) dalam kategori *satisfactory*, 13 soal (soal nomor 4, 5, 8, 13, 14, 15, 16, 17, 31, 32, 33, 35, 38) *good*, 10 soal (soal nomor 6, 10, 18, 19, 22, 24, 27, 34, 36, 40) *excellent* dan 12 soal (soal nomor 3, 9, 11, 12, 23, 25, 26, 28, 29, 30, 37, 39) yang ditolak. Ditinjau dari efektifitas distraktor, dari soal paket A terdapat 34 soal yang distraktornya sudah berfungsi dengan baik, dan 6 soal yang distraktornya belum berfungsi dengan baik. Sedangkan untuk paket soal B, 34 soal distraktornya sudah memenuhi kriteria, 5 soal terdapat 1 distraktor yang belum berfungsi dan 1 soal dengan 2 distraktor yang belum berfungsi.

Soal yang telah diujicobakan dan dianalisis selanjutnya dipilih untuk digunakan pada uji coba II. Pemilihan soal didasarkan pada hasil analisis ITEMAN yang meliputi peninjauan taraf kesukaran item, daya pembeda dan efektifitas distraktor. Dari 80 soal dari kedua paket, dipilih 50 soal untuk uji coba II. Pengambilan keputusan didasarkan pada kriteria sebagai berikut : (1) item soal diterima apabila karakteristik item soal memenuhi semua kriteria. Item soal terlalu sukar atau terlalu mudah tetapi memiliki daya pembeda dan distribusi pengecoh yang memenuhi kriteria maka butir soal tersebut dapat diterima atau dipilih, (2) item soal direvisi, apabila salah satu atau lebih dari ketiga kriteria karakteristik item soal tidak memenuhi syarat, (3) item soal ditolak, jika item soal memiliki karakteristik yang tidak memenuhi semua kriteria (Yusliana, 2000:187).

Selanjutnya dipilih 50 soal yang mewakili setiap indikator untuk digunakan dalam uji coba II sebagai berikut 3A, 2B, 4A, 4B, 5A, 6B, 8B, 11A, 10B, 12A, 13B, 16B, 16A, 17B, 18B, 18A, 19A, 20A, 21A, 22B, 27B, 23A, 24B, 26A, 31B, 30A, 33B, 34A, 35B, 35A, 36B, 39A, 37A, 38B, 40B, dan soal yang dipakai dengan melalui tahap revisi antara lain soal 5B, 2A, 3B, 7B, 7A, 10A, 29A, 29B.

Uji Coba tahap II merupakan suatu tahap akhir penyusunan instrumen tes. Tahap uji coba II dilakukan untuk mengetahui kualitas dari instrumen yang telah disusun. Uji coba dilakukan pada Rabu, 30 Mei 2012 di kelas VIII G, VIII F dan VIII A. Dari uji coba II diperoleh 22 %

soal kategori mudah yakni soal nomor 1, 4, 12, 13, 18, 20, 29, 35, 39, 44 dan 45. Soal termasuk dalam kategori mudah mengindikasikan bahwa soal tersebut mampu dikerjakan oleh peserta tes kelas atas dan peserta tes kelas bawah. Soal yang tergolong sedang sebanyak 54% yaitu soal nomor 2, 3, 5, 7, 8, 9, 10, 11, 14, 15, 16, 21, 23, 25, 26, 27, 28, 30, 31, 34, 36, 38, 41, 42, 43, 47 dan 49. Soal termasuk dalam kategori sedang menunjukkan bahwa soal tersebut dapat dijawab oleh kelompok atas dan kelompok bawah dengan proporsi yang tepat artinya peserta tes kelas atas yang mampu menjawab jumlahnya lebih banyak jika dibandingkan dengan peserta tes yang mampu menjawab dari kelas bawah. Sedangkan soal yang dapat dikategorikan sulit adalah soal nomor 4, 6, 16, dan 18 dengan prosentase 24% yakni soal nomor 6, 17, 19, 22, 24, 32, 33, 37, 40, 46, 48 dan 50. 24% soal yang memiliki tingkat kesukaran tinggi menunjukkan bahwa untuk menjawab soal tersebut diperlukan tingkat kemampuan yang tinggi sehingga hanya peserta tes dari kelas atas yang mampu menjawab soal dengan tepat.

Analisis mengenai daya pembeda menunjukkan bahwa 18% memiliki daya beda *poor*, yakni soal nomor 4, 15, 35, 36, 37, 38, 40, 44 dan 45. Soal yang daya pembedanya ditolak sebanyak 8%, yaitu soal 6, 22, 24 dan 49. Daya pembeda soal *poor* dan ditolak merupakan kategori daya pembeda yang jelek, soal dikatakan memiliki daya pembeda jelek artinya soal – soal tersebut tidak mampu membedakan peserta tes kelas atas dan kelas bawah. Soal yang termasuk kategori *poor* dan ditolak sebaiknya direvisi terlebih dahulu dengan mengkaji ulang kemungkinan penyebab rendahnya daya pembeda soal tersebut. Hal – hal yang dapat menjadi sebab rendahnya nilai daya pembeda misalnya, pilihan jawaban yang membingungkan, pernyataan soal yang kurang jelas atau mungkin peserta tes yang tidak memahami konsep sehingga menganggap soal terlalu sulit sehingga siswa memilih jawaban secara acak.

Soal yang memiliki daya pembeda *satisfactory*, *good*, dan *excellent* secara berturut – turut berjumlah 28%, 36% dan 10%. Soal yang memiliki daya pembeda *satisfactory* yakni soal nomor 1, 3, 7, 14, 17, 19, 20, 29, 39, 41, 42, 43, 47 dan 48. Soal yang memiliki daya pembeda *good* sebanyak 36% yakni 2, 5, 8, 9, 10, 11, 12, 13, 16, 18, 21, 23, 28, 30, 31, 32, 33 dan 46. Sedangkan 10% soal yang termasuk *excellent* yakni soal nomor 25, 26, 27, 34 dan 50. Soal yang termasuk dalam kategori *satisfactory*, *good*, dan *excellent* merupakan soal yang baik, artinya soal tersebut mampu membedakan peserta tes kelas atas dan peserta tes yang berasal dari kelas bawah. Soal yang baik selanjutnya dapat dimasukkan dalam bank soal dan siap untuk diimplementasikan.

Dari analisis ITEMAN mengenai efektifitas distraktor diperoleh 16 soal yang distraktornya belum berfungsi dengan baik, yakni soal nomor 4, 11, 12, 13, 15, 16, 20, 27, 31, 39, 40, 41, 43, 44, 46, 48. Pilihan jawaban yang distraktornya kurang efektif atau tidak berfungsi dengan baik sebaiknya direvisi jika akan digunakan.

## KESIMPULAN

Tes yang telah disusun yaitu “Tes Formatif Fisika SMP Materi Cahaya, tahun Pelajaran 2011/2012”. Tes berupa tes objektif bentuk pilihan ganda dengan empat pilihan jawaban yang berjumlah 50 soal. Alokasi waktu pengerjaan tes selama 90 menit. Tes yang telah disusun dilengkapi dengan petunjuk pengerjaan dan kunci jawaban. Adapun karakteristik tes yang telah disusun adalah

### a. Soal paket A

- 1) Dari segi taraf kesukaran terdapat 25% kategori mudah, 60% kategori sedang dan 15% kategori sukar.



- 2) Daya beda sangat baik (*excellent*) sebanyak 30%, baik (*good*) sebanyak 32,5%, cukup (*satisfactory*) sebanyak 12,5%, rendah (*poor*) sebanyak 12,5% dan ditolak sebanyak 12,5%.
  - 3) Aspek keefektifan distraktor diperoleh hasil 85% soal empat distraktor berfungsi dan 15% soal 3 distraktor berfungsi, 0% soal dua distraktor berfungsi dan 0% soal satu distraktor berfungsi.
- b. Soal paket B
- 1) Dari segi taraf kesukaran terdapat 10% kategori mudah, 57,5% kategori sedang dan 32,5% kategori sukar.
  - 2) Daya beda sangat baik (*excellent*) sebanyak 25%, baik (*good*) sebanyak 32,5%, cukup (*satisfactory*) sebanyak 10%, rendah (*poor*) sebanyak 2,5% dan ditolak sebanyak 30%.
  - 3) Dari aspek keefektifan distraktor diperoleh hasil 85% soal empat distraktor berfungsi dan 12,5% soal 3 distraktor berfungsi, 2,5% soal dua distraktor berfungsi dan 0% soal satu distraktor berfungsi.
- c. Instrumen tes akhir
- 1) Dari segi taraf kesukaran terdapat 22% kategori mudah, 54% kategori sedang dan 24% kategori sukar.
  - 2) Daya beda sangat baik (*excellent*) sebanyak 10%, baik (*good*) sebanyak 36%, cukup (*satisfactory*) sebanyak 28%, rendah (*poor*) sebanyak 18% dan ditolak sebanyak 8%.
  - 3) Dari aspek keefektifan distraktor diperoleh hasil 68% soal empat distraktor berfungsi dan 28% soal 3 distraktor berfungsi, 4% soal dua distraktor berfungsi dan 0% soal satu distraktor berfungsi.
- d. Dari ketiga paket tes, tes yang memiliki kualitas dan karakteristik terbaik adalah tes terakhir.

## DAFTAR PUSTAKA

1. Anas Sudijono. 2005. *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja Grafindo Persada.
2. Azwar, Saifuddin. 2007. *Metode Penelitian*. Yogyakarta : Pustaka Pelajar
3. Azwar, Saifuddin. 1997. *Reliabilitas dan Validitas*. Yogyakarta : Pustaka Pelajar
4. Aminah, N.S. 2002. *Dasar- Dasar Evaluasi Penelitian Pembelajaran*. Surakarta : UNS Press
5. Aminah, N.S. 2012. *Dasar- Dasar Pengukuran dan Statistik pada Pembelajaran Fisika*. Surakarta : UNS Press
6. Djemari Mardapi. 2008. *Teknik Penyusunan Instrumen Tes dan Non Tes*. Yogyakarta: Mitra Cendekia Press.
7. Ekawati, Estina 2011. *Pengembangan Instrumen Penilaian Pembelajaran Matematika Dasar*. Yogyakarta : Kementerian Pendidikan Nasional
8. Elvin Yusliana. (2010). Pemanfaatan Telaah Kualitatif dan Kuantitatif Sebagai Prasyarat Penyelesaian Item Tes Pilihan Ganda Dalam Penyelesaian tugas Akhir (Skripsi) di LPTK. *Akademika II* (2), 181 - 187
9. Mundilarto. 2010. *Penilaian Hasil Belajar Fisika*. Jogjakarta : P2IS FMIPA UNY
10. Suharsimi Arikunto. 2005. *Dasar- Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara. *yang Mengacu pada Standar Penilaian di SMP*. Yogyakarta ; PPPPTK Matematika
11. Supardi. (2008). Pengaruh Bentuk Tes Formatif dan Kecerdasan Emosional terhadap Hasil Belajar Matematika. *Ilmiah Faktor Eksakta I* (2), 18 – 32
12. Yamtinah, Sri. 2009. *Evaluasi Proses dan Hasil Belajar Kimia*. UNS Press : Surakarta